



Antigenic properties of *Fingoldia magna* protein L and Type IV Pilin (PilA) for in-silico multi epitope peptide vaccine designing

Hossein Javid¹

¹Assistant Professor, Department of Genetics, Fars Academic Center for Education, Culture and Research, ACECR, Shiraz, Iran.

Abstract

Background & Objectives: *Fingoldia magna* is a potential opportunistic pathogen for humans. *F. magna* as the most frequent pathogenic species of Gram positive anaerobic cocci accounts for up to 5-12% of all anaerobic infections. *F. magna* possess Protein L super antigen and Type IV Pilin (PilA) that due to trigger intense immune responses are invaluable for designing multi epitope peptide vaccines in current study.

Materials & Methods: In this study, immunoinformatics tools were used to predict B and T cell epitopes of Protein L and Type IV Pilin (PilA). The epitopes were evaluated for antigenicity, allergenicity and binding energy to appropriate DRB3*01:01, DRB1*03:01 and DRB1*15:01 HLA alleles and then were fused together by GPGPG and EAAAK spacers. *Vibrio cholera* Toxin B Subunit was introduced at N-terminus of the constructed vaccine as adjuvant, and with an eye on further identification and purification, a 6×HisTag was introduced at C-terminus. Codon optimization performed for further expression in *Escherichia coli* host. The amino acid sequence of the multi epitope peptide vaccine used for 3D structure prediction and refinement. Then structural evaluation via ramachandran plot analysis performed. Physicochemical properties and solubility of the constructed vaccine was also studied.

Results: Results showed the selected epitopes with high antigenicity and no allergenicity. These epitopes manifest high affinity toward recommended HLA alleles. The predicted 3D model of constructed vaccine showed high stability, solubility and half-life for expression in *E. coli* host.

Conclusion: In this study, Protein L and Type IV Pilin (PilA) used for in-silico designing an effective vaccine against *F. magna*.

Keywords: *Fingoldia magna*, protein L PilA, multi epitope peptide vaccine, HLA.

Received: March 2020 **Accepted:** April 2020

Correspondence to: Hossein Javid

Tel: +98 7132335010

E-mail: h.javid@royaninstitute.org

Journal of Microbial World 2020, 13(1): 87-98.



Copyright © 2019, This article is published in Journal of Microbial World as an open-access article distributed under the terms of the Creative Commons Attribution License. Non-commercial, unrestricted use, distribution, and reproduction of this article is permitted in any medium, provided the original work is properly cited.



ویژگی های آنتی ژنی پروتئین L و پیلین تایپ ۴ (Pila) باکتری *فاینگولدیا ماگنا* (*Finogoldia magna*) به منظور طراحی درون رایانه‌ای واکسن پپتیدی چند اپی توپی

حسین جاوید^{۱*}

استاد یار، جهاد دانشگاهی فارس، گروه پژوهشی ژنتیک

چکیده

سابقه و هدف: *فاینگولدیا ماگنا* یک پاتوژن فرصت طلب انسانی است و شایع ترین گونه بیماری زای کوکسی های بی هوازی گرم مثبت می باشد. این باکتری عامل ایجاد ۵ تا ۱۲ درصد تمام عفونت های بی هوازی است. در این پژوهش سوپر آنتی ژن پروتئین L و پیلین نوع ۴ (Pila) آن به دلیل ایجاد پاسخ های ایمنی قوی و اهمیت در طراحی واکسن های پپتیدی چند اپی توپی مورد استفاده قرار گرفتند.

مواد و روش ها: ابزارهای ایمونولوژی محاسباتی برای پیش بینی اپی توپ ها به کار برده شدند. اپی توپ ها از نظر ویژگی آنتی ژنی و آلرژی‌زایی و انرژی اتصال به آلل های HLA DRB3*01:01، HLA DRB1*03:01 و HLA DRB1*15:01 بررسی شدند و سپس به وسیله توالی های GPGPG و EAAAK به هم متصل شدند. زیر واحد B سم باکتری و *بیبروکلا* به عنوان اجوانت به انتهای آمینی و توالی HisTag_{6x} به منظور خالص سازی و شناسایی پروتئین، به انتهای کربوکسیلی اضافه گردید. همچنین بهینه سازی کدون ها به منظور بیان در *شریشیا کلی* انجام شد. سپس با تهیه ساختار 3D واکسن و انجام اصلاحات ساختاری، به وسیله تحلیل نمودار رامانچاندرا ارزیابی ساختار صورت گرفت. ویژگی های فیزیکی-شیمیایی واکسن، انحلال پذیری و پایداری آن نیز بررسی گردید.

یافته ها: اپی توپ های انتخابی دارای ویژگی آنتی ژنیستی بالا و بدون آلرژی‌زایی بودند و همچنین دارای قدرت اتصال مناسب به آلل های HLA پیشنهادی داشتند. ساختار سه بعدی واکسن پایداری، انحلال پذیری و نیمه عمر بالایی در میزبان بیانی *شریشیا کلی* نشان داد.

نتیجه گیری: در این مطالعه، پروتئین L و Pila باکتری *فاینگولدیا ماگنا* با موفقیت در طراحی درون رایانه ای واکسن مورد استفاده قرار گرفت.

واژگان کلیدی: *Finogoldia magna*، پروتئین L، Pila، واکسن پپتیدی چند اپی توپی، HLA.

دریافت مقاله: اسفند ۹۸ پذیرش برای چاپ: فروردین ۹۹

(* آدرس برای مکاتبه: شیراز، دانشکده علوم پزشکی، جهاد دانشگاهی فارس

تلفن: ۰۷۱۳۲۳۳۵۰۱۰ پست الکترونیک: h.javid@royaninstitute.org



Introduction

Gram positive coccus, *Finegoldia magna*, formerly *Peptostreptococcus magnus*, is a member of strict anaerobic commensal microbiome in skin, oral cavity, gastrointestinal and genourinary tracts and is considered as a potential opportunistic pathogen for humans and animals. *F. magna* is the most frequent and pathogenic species of Gram Positive Anaerobic Cocci (20-40%) and accounts for up to 5-12% of all anaerobic infections (1). Vast clinical infections and various symptoms are thought to be somehow connected to *F. magna* (2), including wound infections, cardiac and pulmonary infections, pericarditis, mediastinitis, necrotizing pneumonia, empyema, soft tissue and musculoskeletal e.g., necrotizing fasciitis, septic arthritis, native and prosthetic joint infections (3) and polymicrobial vaginosis (2), breast abscesses (4) and diabetic feet (5). Fetal monomicrobial bacteremia and Toxic Shock Syndrome (TSS) cases have been reported ethologically connected to *F. magna* (6). *F. magna* has been found in seminal fluid of idiopathic infertile men analyzed by the Next Generation Sequencing (NGS) techniques (7). *F. magna* possesses five virulence factors (Protein L, Type IV Pilin (PilA), PAB, Suf A and FAF) as major parts of its pathogenic arsenal. Among these factors, Protein L (PL) and Type IV Pilin (PilA) (T4P) are studied vastly. PL is approximately 108 kDa and contains 992 amino acid residues with the ability to bind different classes of immunoglobulins through the light chains (8) with the outcome of protecting *F. magna* from being killed by immune system. T4P which helps adherence of *F. magna* to the host tissues via keratinocytes and extracellular matrix enables *F. magna* colonization, biofilm

formation and interaction with the immune system (9). T4P is approximately 14.8 kDa and contains 134 amino acid residues. Immunization against infections is considered as most triumphant medical treatment ever achieved in which, peptide based vaccines are prospective of immunization to trigger intense long lasting protection via recruiting minimal microbial elements and compounds (10). In the current study by using different immunoinformatics platforms, we predict T cell and B cell epitopes of PL super antigen and T4P of *F. magna* ATCC 29328 as model organism. In addition, in-silico, by fusing the predicted epitope peptides via spacers and adjuvant, we designed an ultimate multi epitope peptide vaccine (MEPV) with possibility to prevent infections related to *F. magna*.

Materials and methods

Retrieval of PL and P4T sequences and assessment of antigenic properties

FASTA format files of amino acid sequences of PL (Q51918) and T4P (B0RZY1) of *F. magna* ATCC 29328 and *Vibrio cholera* Toxin B Subunit (CTB; Q7X2D2) were obtained from the Uniprot database at <http://www.uniprot.org>. VaxiJen server at <http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html> (11) was used to predict antigenic-immunogenic attributes of these proteins where the threshold value was set at ≥ 0.4 to increase prediction precision and decreasing unfavorable positively false results.

Prediction and selection of linear B cell epitopes

Prediction of PL and T4P linear B cell epitopes was performed by the immune epitope database and analysis resource (IEDB) server at <http://tools.immuneepitope.org/bcell/> (12) via

Bepipred Linear Epitope Prediction methods, which utilizes a combination of a hidden markov model and a propensity scale method with default threshold value of 0.35. VaxiJen server was used for antigenic properties assessment and AlgPred server at <http://crdd.osdd.net/raghava/algpred/> (13) was used to predict allergenicity of the predicted epitopes.

Prediction and selection of T cell epitopes

The major histocompatibility complex class II (MHC II) binding predictions tool on the IEDB server at <http://tools.iedb.org/mhcii/> (14) was used for T cell epitopes prediction by using the Consensus, Combinatorial library, NN-align (netMHCII-2.2), SMM-align (netMHCII-1.1), Sturniolo, and NetMHCIIpan prediction methods. For this, the 7-allele HLA reference was selected as is recommended (15): HLA-DRB1*03:01, HLA-DRB1*07:01, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01 and HLA-DRB5*01:01 alleles. The percentile ranking cut off ≤ 0.5 was used, where the lower rank manifests the better MHC-II binder. VaxiJen and AlgPred servers were used as mentioned above.

Molecular docking of selected T cell epitopes with HLA alleles

For Molecular docking a high resolution global peptide-protein docking server, PIPER-FlexPepDock, at <http://piperfpd.furmanlab.cs.huji.ac.il> (16) was used. To perform docking, 3D structures (.pdb files) of HLA-DRB1*03:01, HLA-DRB3*01:01 and HLA-DRB1*15:01 retrieved from RCSB protein data bank at <http://www.rcsb.org>. The prepared .pdb files along with selected T cell epitopes were submitted to PIPER-FlexPepDock web server for docking

presses.

Construction and engineering of MEPV, secondary and tertiary structure prediction

The repeated sequences of B cell and T cell epitopes were omitted and then the selected epitopes were fused together via a G and P rich spacer (GPGPG). CTB was used as adjuvant and fused to the amino terminus of the construct via EAAAK linkers. In addition, for further convenient purification and identification, a 6×HisTag was introduced at carboxyl terminus of the final vaccine construct. The secondary structure of the constructed MEPV was predicted based on primary amino acid sequences using PSIPRED web server at <http://bioinf.cs.ucl.ac.uk/psipred/> (17). The amino acid sequences of MEPV, then was uploaded on the I-TASSER server at <http://zhanglab.ccmb.med.umich.edu/I-TASSER/> for 3D structure prediction (18). For high resolution protein structure refinement, the best produced models (based on C and TM scores and RMSD) were uploaded into Modrefiner server at <https://zhanglab.ccmb.med.umich.edu/ModRefiner/> (19). After protein structural refinement, the quality and validation assessment were performed by using Ramachandran plot analysis in RAMPAGE server at <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php> (20).

Assessment of physicochemical parameters of the constructed MEPV

Assessment of various physicochemical parameters and properties of the constructed MEPV was performed using the ProtParam server at <https://web.expasy.org/protparam/> (21). The major assessed parameters include theoretical isoelectric point (pI), grand average of hydropathicity (GRAVY), aliphatic index,

instability index, in-vitro and in-vivo stability. Moreover, the solubility of constructed MEPV was predicted via Protein-Sol online web server available at <https://protein-sol.manchester.ac.uk/> (22).

Optimization of MEPV for overexpression in E. coli host

The OPTIMIZER web server at <http://genomes.urv.es/optimizer/> (23) was used for optimization of constructed MEPV, based on amino acid sequence, for further overexpression in *E. coli* host.

Results

Prediction and selection of linear B cell epitopes

The full length of PL and T4P amino acid sequences were submitted to the B cell epitope prediction tool with default threshold value

0.35 at IEDB server (BepiPred) to predict varying residual length epitopes (Table 1). Three bottle necks applied for selecting most suitable epitopes among IEDB server outputs. Firstly, regarding to the length, only B cell epitopes ≥ 10 mer were selected. Secondly, these selected epitopes were checked for antigenicity at VaxiJen server. Thirdly, epitopes were submitted to AlgPred server for allergenicity prediction. The epitopes did not contain any experimentally proven IgE epitopes and no hit found in blast search on allergen representative peptides (ARPs). Eventually, 16 and 3 B cell epitopes were selected from PL and T4P, respectively, and used for construction of MEPV.

Prediction and selection of T cell epitopes

On the IEDB server, the T cell Epitope

Table 1. The PL and T4P linear B cell epitope peptides.

Protein	Peptide	Length (mer)	Position		VaxiJen Score	Allergenicity
			Start	End		
PL	HKKDEEPGKKPGEDKKPEDKKPGEDKKPEDKKPGEDKKPGKTDKDSPNKKKKAKLPKAGSEA	67	902	968	1.3958	The peptides do not contain experimentally proven IgE epitopes. No hit found in Blast search on Allergen Representative Peptides (ARPs)
	ADEEPPGEDTPEVQEGYATYEEAEAAAKEALKEDKVN	38	653	690	0.8397	
	HAGEETPELKDGYATYEEAEAAAKEALKNDVNN	34	599		0.8916	
	MERKTAEEKAAEKLAATAKETAKK	23	119	141	1.0222	
	GKTQTAEFKGTFFEEATAEAY	20	488	507	0.8942	
	GKTQTAEFKGTFFEEATAK	18	340	357	1.0293	
	DKEDEQPGEEPGENPGIT	18	710	727	1.4060	
	KENGEYTADLEDGGNT	16	367	382	0.9465	
	AGKETPETPEEPKEEV	16	388	403	0.5106	
			462	477		
	IQTAEFKGTFFEEATAK	16	416	431	0.8916	
	LSEKETPEPEEEVT	14	245	258	0.6653	
	NGEYTADLEDGGNT	14	443	456	0.9498	
	ALKKDNGEYTVDV	13	292	304	0.5461	
	GKKEKPEEPKEEV	13	317	329	0.7853	
	AGKEQPGENPGIT	13	536	548	1.5667	
AEKPGENPGIT	11	779	789	1.3621		
		841	851			
T4P	KSVPKVPASNENLK	15	84	98	0.4664	
	ESYRQDNDGKLP	13	60	72	0.9966	
	YKNQKEKAAI	10	36	45	0.6089	

The PL and T4P linear B cell epitope peptides were predicted by IEDB server (BepiPred). Predicted epitopes were selected for MEPV construction based on their length (≥ 10), antigenicity and allergenicity.

Prediction Tool uses six different methods namely consensus, combinatorial library, NN-align, SMM-align, sturniolo and NetMHCIIpan methods. Totally, 6847 and 841 binding sites (15 mer) were predicted on PL and T4P, respectively. The predicted epitopes arranged by percentile ranking (cut off ≤ 0.5) where lower value shows better binding. Epitopes then checked by VaxiJen and AlgPred servers as mentioned before. Eventually, five MHC class II binding epitopes were selected for each PL and T4P (Table 2). For PL, ELKEAGITSDLYFSL, KEAGITSDLYFSLIN, TIKVNLIFADGKTQT and IKVNLIFADGKTQTA epitopes bind prominently to DRB3*01:01 and NPGITIDEWLLKNAK binds to DRB1*03:01, while all 5 predicted T4P epitopes bind to DRB1*15:01. All these 10 epitopes and 3 MHC class II molecules were used for molecular docking.

Molecular docking of selected T cell epitopes with selected HLA alleles

The Flexible protein-peptide docking and

scoring using the Internal Coordinate Mechanics (ICM) were performed using PIPER-FlexPepDock web server. For each docking a total number of 10 models were produced by PIPER-FlexPepDock web server and only the first model with the lowest binding energy was selected. The results revealed that the binding energy between ELKEAGITSDLYFSL, KEAGITSDLYFSLIN, TIKVNLIFADGKTQT and IKVNLIFADGKTQTA of PL T cell epitopes sequences and DRB3*01:01 HLA allele were -756.3, -775.4, -768.5 and -792 kcal/mol, respectively. The binding energy between NPGITIDEWLLKNAK and DRB1*03:01 HLA allele was -375.7 kcal/mol. The binding energy between LVVLAILAILIAIAV, LLELLVVLAILAILI, LELLVVLAILAILIA, ELLVVLAILAILIAI and LLVVLAILAILIAIA of T4P T cell epitopes sequences and DRB1*15:01 HLA allele were -578.1, -613.7, -601.7, -599.1 and -601.4 kcal/mol, respectively. The results showed that all T cell epitopes acceptably bind to their corresponding HLA alleles, whereas ≤ -32 kcal/mol typically is considered as good

Table 2. The PL and T4P T cell epitope peptides.

Protein	Peptide (15 mer)	Position		HLA Allele	Consensus Methods	Percentile Ranking	Comb.lib Core Sequence	Smm Align Core Sequence	Nn Align Core Sequence	Sturniolo Core Sequence	VaxiJen Score	Allergenicity	
		Start	End										
PL	ELKEAGITSDLYFSL	564	578	DRB3*01:01	comb.lib./smm/nn	0.03	ITSDLYFSL	ITSDLYFSL	ITSDLYFSL	-	1.1186	The peptides do not contain experimentally proven IgE epitopes. No hit found in Blast search on Allergen Representative Peptides (ARPs)	
	KEAGITSDLYFSLIN	566	580								0.9307		
	NPGITIDEWLLKNAK	544	558	DRB1*03:01	smm/nn/sturniolo	0.54	-	ITIDEWLLK	ITIDEWLLK	ITIDEWLLK	1.1014		
		723	737										
	TIKVNLIFADGKTQT	785	799	DRB3*01:01	comb.lib./smm/nn	0.6	IKVNLIFAD	IKVNLIFAD	IFADGKTQT	-	0.9390		
		847	861										
	IKVNLIFADGKTQTA	330	344	DRB1*15:01	smm/nn/sturniolo	0.26	-	VVLAILAIL	VVLAILAIL	VVLAILAIL	0.9264		
		478	492										
	T4P	LVVLAILAILIAIAV	331	345	DRB1*15:01	smm/nn/sturniolo	0.23	-	VLAILAILI	ILAILIAIA	LLVVLAILAILIA		1.0888
		LLELLVVLAILAILI	19	33									0.9445
LELLVVLAILAILIA		15	29	0.9780									
ELLVVLAILAILIAI		16	30	0.9678									
LLVVLAILAILIAIA		17	31	1.0235									

The PL and T4P T cell epitope peptides were predicted by IEDB server (MHC II binding predictions tool). The predicted epitopes were selected based on their binding energy to related HLA, antigenicity and allergenicity. The selected epitope peptides were used for MEPV construction.

binding in ICM method (24).

Construction of MEPV, secondary and tertiary structure prediction

Fig. 1a shows the final construction of MEPV. Totally 16 and 3 sequences of PL and T4P B cell epitopes, respectively, and 10 T cell epitopes (5 epitopes from each PL and T4P), were fused together by GPGPG spacers. A 6×HisTag was introduced at carboxyl terminus with the same spacer for further purification with Immobilized Metal Affinity Chromatography (IMAC) and detection of MEPV recombinant vaccine via Western Blot technique by using 6×HisTag specific antibodies or probes. CTB was introduced at amino terminus of the construct as adjuvant with EAAAK spacer. Fig. 1b shows the position of residues contribute to strands, helix and coil structures prepared by PSIPRED web server and Fig. 1c shows the residues positions regarding to small non-polar, polar, hydrophobic and aromatic nature. Fig. 1d shows PSIPRED graphical result of MEPV secondary structure prediction with confidence for each residue in a color saturation manner. The results showed that approximately 26.97%, 23.33% and 16.33% of the residues contribute to helix, coil and strand structures, respectively, and 33.37% contribute to turn and bridge structures. Among the five best generated 3D models of MEPV, the highest C-score was -0.13 that is in acceptable range. The TM-score and estimated RMSD are 0.70 ± 0.12 and $8.7 \pm 4.5 \text{ \AA}$, respectively (Fig. 2a). Based on Ramachandran plot analysis, in MEPV initial model, 65.7%, 19.9% and 14.5% of residues were positioned at favored, allowed and outlier regions, respectively (Fig. 2b). Whereas, for MEPV refined model values improved to 79.6%, 14.7% and 5.7% (Fig. 2c).

Assessment of physicochemical parameters of MEPV

The final MEPV construct (C₃₆₀₄H₅₇₁₇N₉₆₁O₁₁₈₁S₇) with 797 amino acid residues is estimated to have 81.66 kDa molecular weight and theoretical isoelectric point value of 4.8, indicating negative nature of the protein with 126 negatively charged residues (E: 94 and D: 34) and 84 positively charged residues (R: 5 and K: 79). The grand average of hydropathicity (GRAVY) predicted to be -0.682 which indicates MEPV to be hydrophilic. The half-life predicted to be 30h in mammalian reticulocytes in-vitro, >20h in yeasts and >10h in *E. coli* in-vivo. The instability index of MEPV is calculated to be 21.10 which classifies it as stable. Moreover, MEPV shows aliphatic index of 66.71 as a clue for its thermos ability. The result obtained from Protein-Sol sever shows the predicted scaled solubility (0-1) of MEPV to be 0.777, where the value greater than 0.45 (the average of soluble *E. coli* proteins) is predicted to have a higher solubility than that of the average soluble *E. coli* proteins.

Codon optimization

The amino acid sequence of MEPV was submitted to OPTIMIZER web server for codon optimization in *E. coli* strain K-12 with One AA-one codon criteria. The result showed the CAI-value and GC content of optimized MEPV improved to 1.00 and 57%, respectively, while CAI-value and GC content of 0.88 and 59.6%, respectively, were calculated for un-optimized MEPV. The CAI-value > 0.8 is normally considered as good for high level expression in *E. coli* host cells. The suitable CG content for expression in *E. coli* host is between 30% and 70% (25).

Discussion

In recent years, in-silico progress in prediction

and understanding of epitopes recognition at molecular level have enlightened the path for multi epitope peptide vaccine development

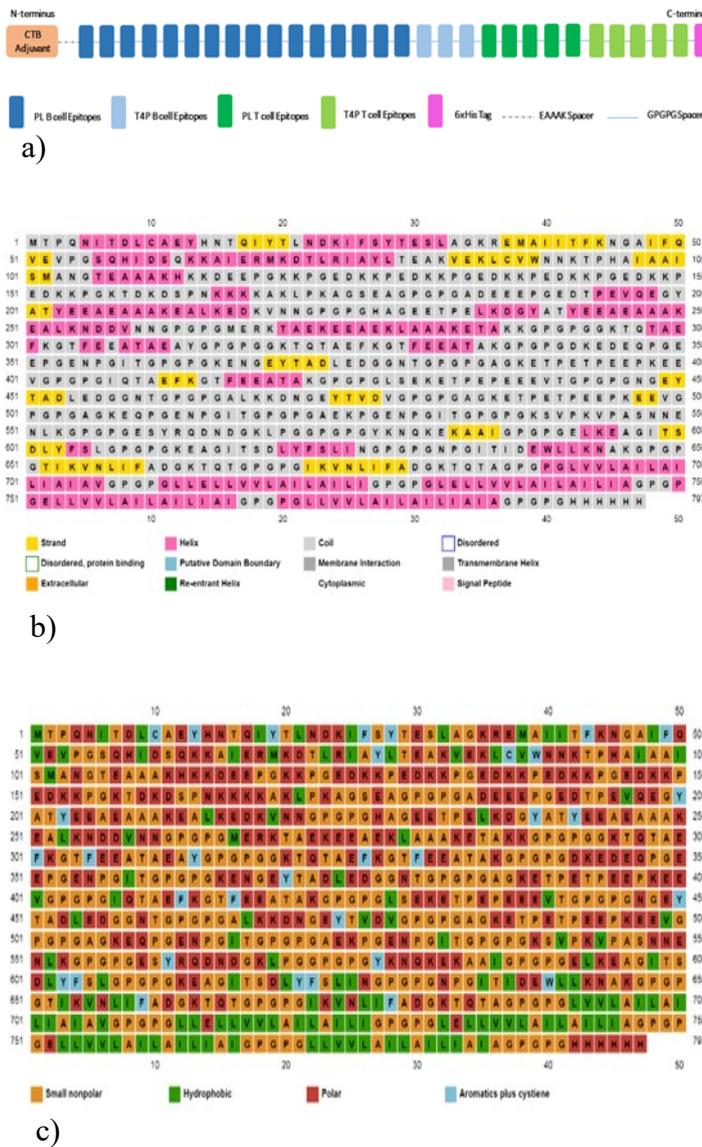
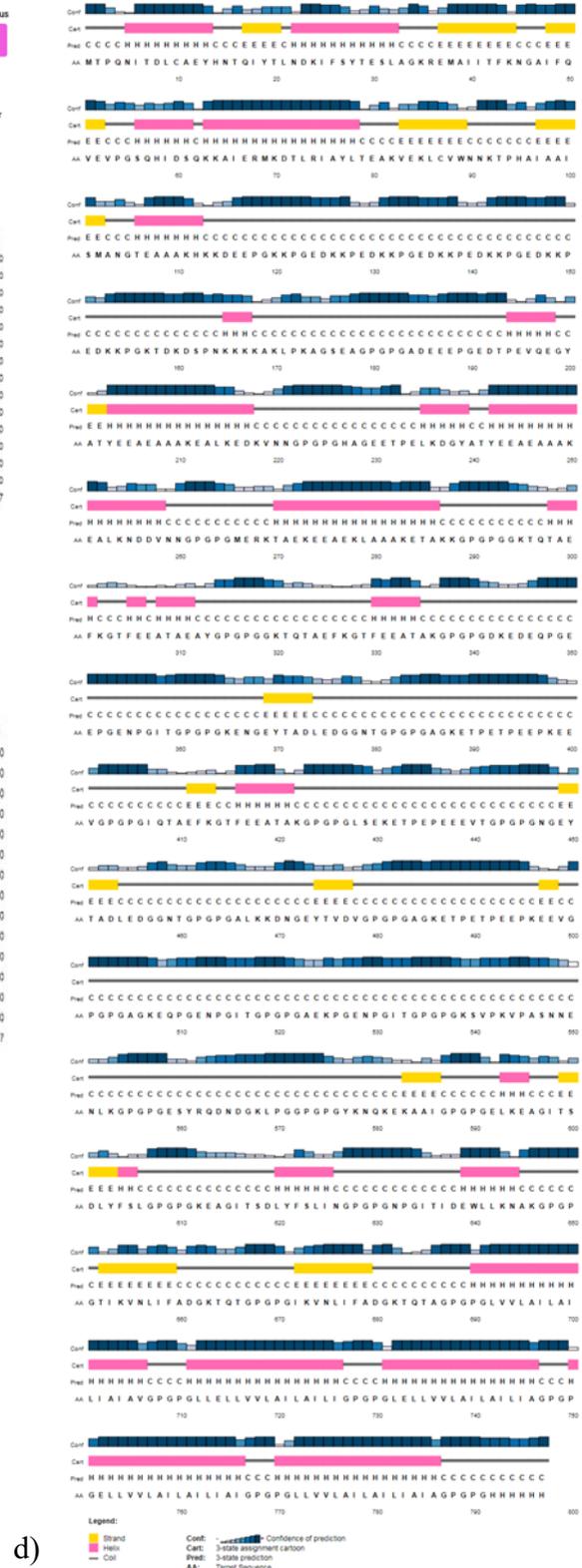


Fig. 1. Construction of multi epitope peptide vaccine and secondary structure. a) the final construction of MEPV results from fusion of B and T cells epitopes, CTB adjuvant and 6×HisTag by GPGPG and EAAAK spacers, b) position and frequency of residues contribute to strand, helix and coil structures and c) regarding to their nature, d) the PSIPRED graphical result of MEPV secondary structure prediction with confidence for each residue in a color saturation manner.



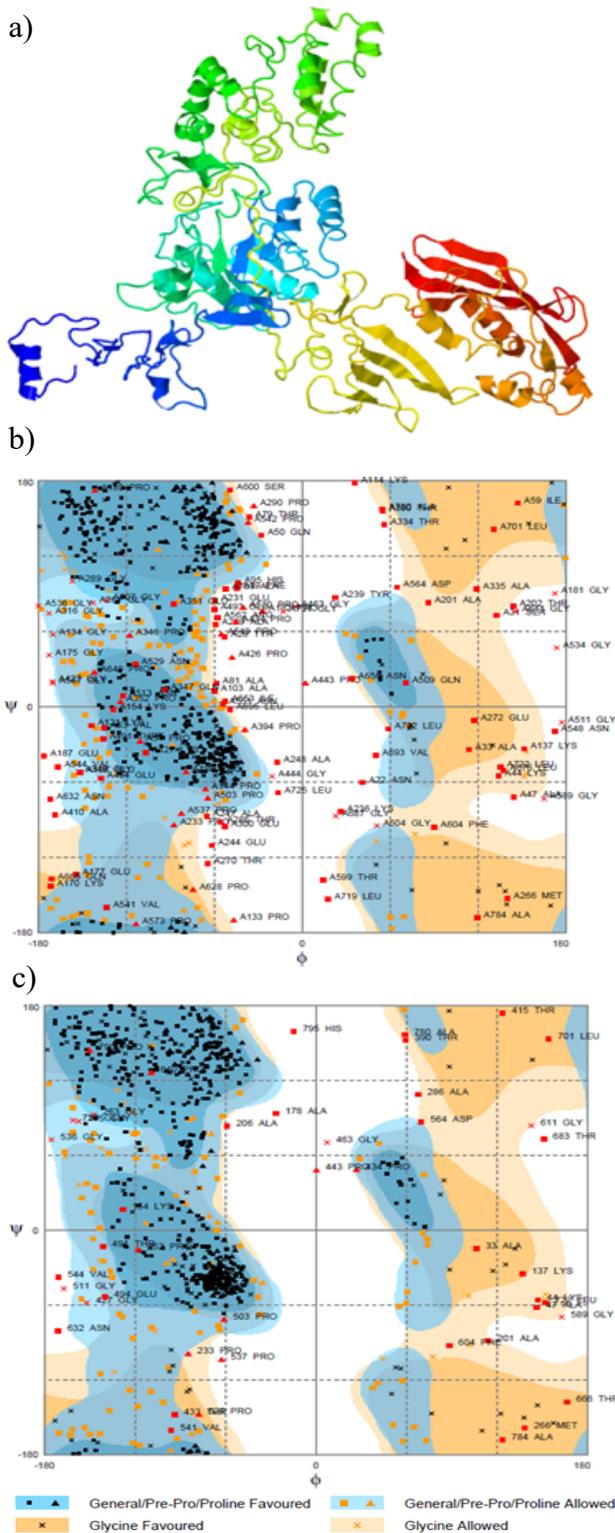


Fig. 2. 3D structure prediction of MEPV. a) the final 3D model of MEPV, b) ramachandran plot of initial MEPV model before refinement and c) after refinement.

relies on B cell and T cell epitope prediction. *F. magna* is believed to be one of the most important member of Gram Positive Anaerobic Cocci and cause of severe infections in many parts of body (1). A suitable and universal multicomponent vaccine to reinforce immune response against this bacterium is yet to be developed and is the object of the current work. Our strategy in this study was designed to select best T cell and B cell epitopes of two key virulence factors of *F. magna* by using open source servers. The predicted epitopes were checked both for antigenicity and allergenicity on VaxiJen and AlgPred web servers, respectively. VaxiJen web server developed as the first server alignment-independent prediction of protective antigens and a protein query can be submitted to this server followed by selection of three targets and antigen data bases (11). To assay allergenicity, two salient features of AlgPred server were used: firstly, the mapping of IgE approach in which the query protein sequence will be checked for any known IgE epitopes. The protein will be assign as allergen if high similarity with IgE epitopes be found. This approach has very high specificity. Secondly, a blast search on ARPs was performed. In this approach the query protein sequence is searched against almost 2890 known ARPs. If there be any hit, the query protein is assign as allergen. The accuracy of this method is very high with high sensitivity and specificity (13). Neither of the epitopes and final MEPV showed allergenicity, which strengthen the potential of MEPV as vaccine candidate. Raising suitable immune responses strongly depends on the interaction between epitopes and related receptors. All the predicted T cell epitopes were used in docking with recommended HLA alleles to determine the

best and lowest binding energy by PIPER-FlexPepDock web server. This server streamlines rosetta fragment picker available at <https://www.rosettacommons.org> to generate precise peptide fragment ensembles. The server utilizes PIPER docking algorithm for fragment-receptor rigid-body docking and rosetta FlexPepDock for flexible full atom refinement of docked models. Moreover, rosetta cluster exerted for clustering of generated models. The best T cell epitopes with lowest binding energy were selected for further processing. The selected B cell and T cell epitopes were fused together by GPGPG spacers. This spacer in MEPV structure prohibits formation of the junctional epitopes and boosts processing and presenting of epitopes (26), while EAAAK spacer, which used to connect CTB to the rest of MEPV, reduces the interaction between CTB and other parts of MEPV and provides effective separation. CTB is a non-toxic part of Cholera toxin which is a potent adjuvant and mucosal immunostimulant (27). A popular and highly accurate secondary structure prediction server, PSIPRED, was used to predict MEPV secondary structure which showed residues contribute to helices, coils, strand structures, turns and bridges. The prepared MEPV then was submitted to I-TASSER server for 3D structure prediction. Top 5 models for MEPV were generated and arranged according to C-score, TM-score and RMSD. C-score or confidence score value is used for quality prediction of models by I-TASSER server and ranges between 2 and -5, where the higher indicates higher confidence. The TM-score or template modeling is another score to show the correct topology of the model where the score higher than 0.5 is preferable. RMSD along with TM-score are used to confirm reliability and

accuracy of the generated 3D models. The generated 3D .pdb files of MEPV were submitted to Modrefiner server, which used for high-resolution refinement of generated 3D model. The major purpose of using Modrefiner is to refine 3D models closer to their native state and includes repositioning and adjusting of side chains, hydrogen bonds and backbone topology. Both .pdb files of MEPV, before and after refinement, were submitted to RAMPAGE server for quality and validation assessment via Ramachandran Plot Analysis which shows MEPV to be fully stable with 79.6% and 14.7% of residues (94.3%) located at favorite and allowed regions. Finally, the OPTIMIZER server used for codon usage optimization to increase protein expression level. It is possible for users to introduce specific preference optimization tables or select from more than 150 pre-computed tables for prokaryotes in OPTIMIZER server by using 3 approaches e.g., one amino acid-one codon, random or intermediate. Evaluation of effective optimization shows by The codon adaptation index (CAI) which improved from 0.88 to 1.0 for optimized MEPV. physicochemical parameters of MEPV determined by ProtParam server revealed that this protein has acceptable stability and solubility in *E. coli* expression system. Moreover, hydrophobicity along with fused 6×HisTag provide clues for designing further purification methods.

Conclusion

Designing an effective multi epitope peptide vaccines needs advanced knowledge regarding to both B cell and T cell related antigen recognition which could achieve by immunoinformatics tools. To the best of our knowledge in the current study, PL and T4P are used for the first time for in-silico designing of

a multi epitope peptide vaccine with possible protective effects against *F. magna*. Further in-vitro/vivo evaluation of this in-silico designed vaccine is highly suggested to confirm its immunogenicity and protective attributes.

Ethical Consideration

Authors of all ethics including non-plagiarism, Dual publishing has complied with data distortions and data making in this article.

Acknowledgment

The author appreciates Dr. Mazaher Babaie, Dr. Maryam Soroush and Mrs. Sara Senemar at Fars Academic Center for Education, Culture and Research, ACECR, Shiraz, Iran for supporting the conduct of this research.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Arsene C, Saste A, Somiah M, Mestrovich J, Berger G. A case of septic arthritis of the wrist due to *Finegoldia magna*. *Case Rep Infect Dis.* .793053 ;2014
2. Murphy EC, Frick IM. Gram-positive anaerobic cocci—commensals and opportunistic pathogens. *FEMS Microbiol Rev* 2013; 37 (4): 520-553.
3. Söderquist B, Björklund S, Hellmark B, Jensen A, Brüggemann H. *Finegoldia magna* isolated from orthopedic joint implant-associated infections. *J Clin Microbiol.* 2017; 55 (11): 3283-3291.
4. Cobo F, Rodríguez-Granger J, Sampedro A, Navarro-Marí JM. Breast abscess due to *Finegoldia magna* in a non-puerperal woman. *Anaerobe.* 2017; 47: 183-184.
5. Citron D, Goldstein E, Merriam C, Lipsky B, Abramson M. Bacteriology of moderate-to-severe diabetic foot infections and in vitro activity of antimicrobial agents. *J Clin Microbiol.* 2007; 45 (9): 2819–2828.
6. Rosenthal ME, Rojzman AD, Frank E. *Finegoldia magna* (formerly *Peptostreptococcus magnus*): An overlooked etiology for toxic shock syndrome? *Med. Hypotheses* 2012; 79 (2): 138-140.
7. Weng S-L, Chiu C-M, Lin F-M, Huang W-C, Liang C. Bacterial communities in semen from men of infertile couples: metagenomic sequencing reveals relationships of seminal microbiota to semen quality. *PLoS ONE* 2014; 9 (10): e110152.
8. Anderson A, Zheng Y, Song D, LaRosa D, Rooijen N, Kierstein G, Kierstein S, Haczku A, Levinsonl A. The b cell super antigen *Finegoldia magna* protein L causes pulmonary inflammation by a mechanism dependent on myd88 but not b cells or immunoglobulins. *Inflamm Res.* 2012; 61 (2): 161–169.
9. Murphy EC, Janulczyk R, Karlsson C, Mörgelin M, Frick IM. Identification of pili on the surface of *Finegoldia magna* - A Gram positive anaerobic cocci. *Anaerobe* 2014; 27: 40-49.
10. Skwarczynski M, Totha I. Peptide-based synthetic vaccines. *Chem Sci.* 2016; 7 (2): 842-854.
11. Doytchinova I, Flower D. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007; 5: 8-4.
12. Larsen J, Lund O, Nielsen M. Improved method for predicting linear B cell epitopes.

- Immunome Res. 2006; 24: 2-2.
13. Saha S, Raghava G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 2006; 34: 202-209.
 14. Vita R, Zarebski L, Greenbaum J, Emami H, Hoof I. The immune epitope database 2.0. *Nucleic Acids Res.* 2009; 38: 854-862.
 15. Paul S, Arlehamn C, Scriba T, Dillon M, Oseroff C. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. *J Immunol Methods.* 2015; 422: 28-34.
 16. Alam N, Goldstein O, Xia B, Porter K, Kozakov D, Schueler-Furman O. High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. *PLoS Comput. Biol.* 2017; 13: e1005905.
 17. Buchan D, Jones D. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* 2019; 47: 402-407.
 18. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* 2015; 43: 174-181.
 19. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J.* 2011; 101: 2525-2534.
 20. Lovell S, Davis I, Arendall B, Bakker P. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins.* 2003; 50 (3): 437–450.
 21. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M. (2005). Protein identification and analysis tools on the ExPASy server. I In: Walker J.M. (eds) *The Proteomics Protocols Handbook*. Humana Press.
 22. Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J. Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics* 2017; 33 (19): 3098-3100.
 23. Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* 2007; 35: 126-131.
 24. Saha CK, Mahbub Hasan M, Hossain S, Jahan A, Azad AK. In silico identification and characterization of common epitope-based peptide vaccine for Nipah and Hendra viruses. *Asian Pac J Trop Med* 2017; 10 (6): 529-538.
 25. Jin Z, Sun T, Xia X, Wei Q, Song Y. Optimized expression, purification of herpes B virus GD protein in *Escherichia coli*, and production of its monoclonal antibodies. *Jundishapur J Microbiol* 2016; 9 (3): e32183.
 26. Saadi M, Karkhah A, Nouri HR. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infect. Genet. Evol.* 2017; 51: 227-234.
 27. Stratmann T. Cholera toxin subunit B as adjuvant-an accelerator in protective immunity and a break in autoimmunity. *Vaccines* 2015; 3 (3): 579-596.